10/8/21

Last time:

Chain Rule = $\frac{\partial f}{\partial t_i} = \frac{\partial f}{\partial x_1}\left(\frac{\partial x_1}{\partial t_i}\right) + \frac{\partial f}{\partial x_2}\left(\frac{\partial x_2}{\partial t_2}\right) + \dots + \frac{\partial f}{\partial x_n}\left(\frac{x_i}{t_i}\right)$

Implicit Function Theorem: Let F be a function w/ $\frac{\partial F}{\partial x_n} \neq 0$ and $\frac{\partial F}{\partial x_i}$ cts. Then on the locus of $F(x_1, x_2, \dots, x_n) = 0$, we have locally $x_n = f(x_1, \dots, x_{n-1})$ & $\frac{\partial f}{\partial x_i} = -\frac{\partial F}{\partial x_i} / \frac{\partial F}{\partial x_n}$.

Proof (IFT Derivative Formula):

Apply a partial derivative to F using chain rule:

$$0 = \frac{\partial F}{\partial x_1}\left(\frac{\partial x_1}{\partial x_i}\right) + \frac{\partial F}{\partial x_2}\left(\frac{\partial x_2}{\partial x_i}\right) + \dots + \frac{\partial F}{\partial x_n}\left(\frac{\partial x_n}{\partial x_i}\right)$$

For $i \neq k$ and $k \neq n$, we have

$\frac{\partial x_k}{\partial x_i} = 0$, Thus we obtain:

$$0 = \frac{\partial F}{\partial x_i}\left(\frac{\partial x_i}{\partial x_i}\right) + \frac{\partial F}{\partial x_n}\left(\frac{\partial x_n}{\partial x_i}\right) = \frac{\partial F}{\partial x_i} + \frac{\partial F}{\partial x_n}\left(\frac{\partial f}{\partial x_i}\right)$$

Solving we obtain $\frac{\partial f}{\partial x_i} = -\frac{\frac{\partial F}{\partial x_i}}{\frac{\partial F}{\partial x_n}}$ $\blacksquare$

Ex: Compute $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ for implicit function $z(x,y)$ given by $x^3+y^3+z^3 = 2xyz-5$

Sol: We want to use IFT.

$$x^3+y^3+z^3 = 2xyz-5 \quad \text{iff} \quad x^3+y^3+z^3-2xyz+5=0$$

Using, $F(x,y,z) = x^3+y^3+z^3-2xyz+5$, we see

$$\frac{\partial F}{\partial x} = 3x^2-2yz, \quad \frac{\partial F}{\partial y} = 3y^2-2xz, \quad \frac{\partial F}{\partial z} = 3z^2-2xy.$$

Hence by IFT: $\frac{\partial z}{\partial x} = -\frac{\partial F}{\partial x}\Big/\frac{\partial F}{\partial z} = -\frac{3x^2-2yz}{3z^2-2xy}$

and $\frac{\partial z}{\partial y} = -\frac{\partial F}{\partial y}\Big/\frac{\partial F}{\partial z} = -\frac{3y^2-2xz}{3z^2-2xy}$ $\square$

Gradient and Optimization

Goal: Optimize functions of several variables by extending the calculus I tricks to multiple variables.

Def: The gradient of functions $f(x_1, x_2, \ldots x_n)$ is:

$$\nabla f = \left\langle \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right\rangle.$$

NB: The gradient can be used to clearly restate some stuff we've seen.

① **Chain Rule:** $\frac{\partial f}{\partial t_i} = \nabla f \left(\frac{\partial \vec{x}}{\partial t_i} \right)$

by chain rule

Why?! $\frac{\partial f}{\partial t_i} = \frac{\partial f}{\partial x_1}\left(\frac{\partial x_1}{\partial t_i}\right) + \frac{\partial f}{\partial x_2}\left(\frac{\partial x_2}{\partial t_i}\right) + \ldots + \frac{\partial f}{\partial x_n}\left(\frac{\partial x_n}{\partial t_i}\right)$

$$= \underbrace{\left\langle \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right\rangle}_{} \cdot \underbrace{\left\langle \frac{\partial x_1}{\partial t_i}, \frac{\partial x_2}{\partial t_i}, \ldots, \frac{\partial x_n}{\partial t_i} \right\rangle}_{}$$

$$= \nabla f \cdot \frac{\partial \vec{x}}{\partial t_i}$$

Claim: Directional derivative can also be expressed using the gradient...

Why?: Recall that the directional derivative of f at $\vec{p}$ in the direction of unit vector $\vec{u}$ is:

$$D_{\vec{u}} f(\vec{p}) = \lim_{h \to 0^+} \frac{f(\vec{p} + h\vec{u}) - f(\vec{p})}{h}$$

Define $g(h) = f(\vec{p} + h\vec{u})$ and notice $g(0) = f(\vec{p})$.

$\therefore D_{\vec{u}} f(\vec{p}) = \lim_{h \to 0^+} \frac{g(h) - g(0)}{h} = g'(0)$. On the other hand,

$$g'(h) = \frac{\partial}{\partial h}\left[f(\vec{p} + h\vec{u})\right] = \frac{\partial}{\partial h}\left[f(p_1 + hu_1, p_2 + hu_2, \dots, p_n + hu_n)\right]$$

Recognise this as a chain rule for $x_i = p_i + hu_i$:

$$g'(h) = \nabla f(\vec{p} + h\vec{u}) \cdot \frac{\partial \vec{x}}{\partial h} = \nabla f(\vec{p} + h\vec{u}) \cdot \langle u_1, u_2, \dots, u_n \rangle \to$$

$$\to = \nabla f(\vec{p} + h\vec{u}) \cdot \vec{u}$$

$\therefore$ we have:

$$g'(0) = \nabla f(\vec{p} + 0\vec{u}) \cdot \vec{u} = \nabla f(\vec{p}) \cdot \vec{u}.$$

Finally we have:

(2) $D_{\vec{u}} f(\vec{p}) = \nabla f(\vec{p}) \cdot \vec{u}.$

Ex: Compute $D_{\vec{u}} f(\vec{p})$ for $f(x,y) = 4x\sqrt{y}$ at $\vec{p} = \langle 1, 4 \rangle$ in direction $\vec{u} = \langle -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \rangle$.

Sol: We know $D_{\vec{u}} f(\vec{p}) = \nabla f(\vec{p}) \cdot \vec{u}$

$$\nabla f(x,y) = \langle 4y^{\frac{1}{2}}, 2xy^{-\frac{1}{2}} \rangle.$$

$$\therefore \nabla f(p) = \langle 4(2), 2(1)(\tfrac{1}{2}) \rangle = \langle 8, 1 \rangle.$$

$$\therefore D_{\vec{u}} f(\vec{p}) = \nabla f(\vec{p}) \cdot \vec{u} = \langle 8, 1 \rangle \cdot \langle -\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}} \rangle \rightarrow$$

$$\rightarrow = -\frac{8}{\sqrt{2}} + \frac{1}{\sqrt{2}} = -\frac{7}{\sqrt{2}} \quad \boxed{\text{R}}$$

Ex: Compute ∇f for $f(x,y,z) = \dfrac{xz}{y+z}$

Sol: $\nabla f = \left\langle \dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial y}, \dfrac{\partial f}{\partial z} \right\rangle$.

$$\frac{\partial f}{\partial x} = \frac{z}{y+z}, \quad \frac{\partial f}{\partial y} = -\frac{xz}{(y+z)^2}, \text{ and}$$

$$\frac{\partial f}{\partial z} = \frac{(y+z)\frac{\partial}{\partial z}[xz] - xz\frac{\partial}{\partial z}[y+z]}{(y+z)^2} \rightarrow$$

$$\rightarrow = \frac{(y+z)(x) - xz(1)}{(y+z)^2} = \frac{xy}{(y+z)^2}$$

$$\therefore \nabla f = \left\langle \frac{z}{y+z}, -\frac{xz}{(y+z)^2}, \frac{xy}{(y+z)^2} \right\rangle$$

Q: How do we optimize the directional derivative?

Think about f at $\vec{p}$ and vary unit vector $\vec{u}$.

$$D_{\vec{u}} f(\vec{p}) = \nabla f(\vec{p}) \cdot \vec{u} = |\nabla f(\vec{p})||\vec{u}| \cos(\theta) \rightarrow$$

computed earlier   geo property of dot

$$\rightarrow = |\nabla f(\vec{p})| \cos\theta$$

$\vec{u}$ unit vector

$\therefore$ maximizing $D_{\vec{u}} f(\vec{p})$ amounts to maximizing $\cos(\theta)$. We know from Calc I that $\cos(\theta)$ is maximized at $\cos(0) = 1$.

$\therefore$ ① The direction of the gradient maximizes directional derivative.

② The magnitude of the gradient ($|\nabla f|$) is the maximum directional derivative at $\vec{p}$.

Ex: Compute direction and max value of $D_{\vec{u}} f(\vec{p})$ for $f(x,y,z) = \frac{xz}{y+z}$ at $\vec{p} = \langle 1,1,-2 \rangle$.

Sol: We already computed $\nabla f = \langle \frac{z}{y+z}, \frac{-xz}{(y+z)^2}, \frac{xy}{(y+z)^2} \rangle$

$\therefore$ at $\vec{p} = \langle 1,1,-2 \rangle$, the dir. derivative is maximized in direction $\nabla f(1,1,-2) = \langle \frac{-2}{(-2)}, -\frac{1(-2)}{(1-2)^2}, \frac{1(1)}{(1-2)^2} \rangle =$

$\rightarrow = \langle 2,2,1 \rangle$.

Furthermore, the max value is:

$|\nabla f(\vec{p})| = |\langle 2,2,1 \rangle| = \sqrt{4+4+1} = 3$ $\boxed{\text{⸬}}$

Def: A function f has...

① a local maximum value at $\vec{p}$ when $f(\vec{p}) \geq f(\vec{x})$ for all $\vec{x}$ near $\vec{p}$.

② a global maximum value at $\vec{p}$ when $f(\vec{p}) \geq f(\vec{x})$ for all $\vec{x} \in \mathrm{dom}(f)$.

(we call $\vec{p}$ the (local/global) maximum point for f).

③ minima (both local and global) are defined similarly. [Just flip inequality]

Recall: $f(x) = x$ has none of these...

Q: How do we guarentee the existence of extrema?

maxima or minima. Where do we look for them?

Def: A point $\vec{p} \in \mathrm{dom}(f)$ is a critical point of f when either $\nabla f(\vec{p})$ does not exist or $\nabla f(\vec{p}) = \vec{0}$.

Prop (Fermat's Extremum Theorem): The local extrema of function f occur only at critical points of f.